

METRIC PROPERTIES OF THE SET OF ORTHOGONAL PROJECTIONS AND THEIR APPLICATIONS TO OPERATOR PERTURBATION THEORY

KONSTANTIN A. MAKAROV AND ALBRECHT SEELMANN

ABSTRACT. We prove that the set of orthogonal projections on a Hilbert space equipped with the length metric is $\frac{\pi}{2}$ -geodesic. As an application, we consider the problem of variation of spectral subspaces for bounded linear self-adjoint operators and obtain a new estimate on the norm of the difference of two spectral projections associated with isolated parts of the spectrum of the perturbed and unperturbed operators, respectively. In particular, recent results by Kostykin, Makarov and Motovilov from [Trans. Amer. Math. Soc., V. 359, No. 1, 77 – 89] and [Proc. Amer. Math. Soc., 131, 3469 – 3476] are sharpened.

1. INTRODUCTION

The main purpose of this paper is to study metric properties of the (noncommutative) space \mathcal{P} of orthogonal projections acting in a separable Hilbert space \mathcal{H} with the emphasis on applications to the spectral perturbation theory. On the metric space (\mathcal{P}, d) , where d is the metric introduced by the norm in the space $\mathcal{L}(\mathcal{H})$ of bounded operators on \mathcal{H} ,

$$d(P, Q) = \|P - Q\|, \quad P, Q \in \mathcal{P},$$

we introduce the length metric ρ , so that the space (\mathcal{P}, ρ) becomes a length space, with the distance ρ between two points defined as the infimum of the lengths of the paths that join them.

One of our principle results regarding the global geometry of the space of projections \mathcal{P} is that the length space (\mathcal{P}, ρ) is $\frac{\pi}{2}$ -geodesic. That means that any two projections $P, Q \in \mathcal{P}$ with $\rho(P, Q) < \frac{\pi}{2}$ can be connected by a geodesic path of length $l = \rho(P, Q)$. Recall that a path $\gamma : [a, b] \rightarrow \mathcal{P}$ is called a geodesic if

$$\rho(\gamma(t), \gamma(s)) = |t - s|, \quad t, s \in [a, b].$$

In particular, we prove that the collection of the open unit balls in (\mathcal{P}, d) coincides with the one of the open balls of radius $\frac{\pi}{2}$ in the length space (\mathcal{P}, ρ) , that is,

$$\|P - Q\| < 1 \quad \text{iff} \quad \rho(P, Q) < \frac{\pi}{2} \quad \text{for} \quad P, Q \in \mathcal{P}.$$

The pairs (P, Q) of orthogonal projections with $\|P - Q\| < 1$ are of special interest. For instance, such P and Q are unitarily equivalent. Moreover, $\text{Ran } Q$ is a graph subspace of a bounded operator $X : \text{Ran } P \rightarrow \text{Ran } P^\perp$ and hence the relative geometry of the subspaces $\text{Ran } P$ and $\text{Ran } Q$ can efficiently be studied by using standard tools of the geometric perturbation theory. The key role in our study of the relative geometry of the graph subspaces $\text{Ran } P$ and $\text{Ran } Q$ with $\rho(P, Q) < \frac{\pi}{2}$ is played by the operator angle Θ , a self-adjoint operator that can be introduced via the operator X by the functional calculus

$$\Theta = \arctan(X^* X)^{1/2}.$$

Using the concept of the operator angle we show that the length metric ρ is locally characterized by the norm of Θ :

$$\rho(P, Q) = \|\Theta\| \quad \text{if} \quad \rho(P, Q) < \frac{\pi}{2}.$$

Using the characterization of the length metric as the infimum of the arc lengths and the well known relation $\|\Theta\| = \arcsin \|P - Q\|$, we prove the following sharp inequality

$$\arcsin \|P - Q\| \leq \int_a^b \|\dot{\gamma}(t)\| dt$$

relating the norm of the difference of orthogonal projections and the arc length of a smooth path $\gamma : [a, b] \rightarrow \mathcal{P}$ joining them.

As the first application of our geometric study of the space \mathcal{P} to the spectral perturbation theory, we consider a smooth self-adjoint path of bounded operators B_t each having two disjoint spectral components. Given that the two families $\{\omega_t\}_{t \in I}$ and $\{\Omega_t\}_{t \in I}$ of spectral components depend upper semicontinuously on the parameter, we prove the following inequality

$$\arcsin(\|P_t - P_0\|) \leq \frac{\pi}{2} \int_0^t \frac{\dot{B}_\tau}{\text{dist}(\omega_\tau, \Omega_\tau)} d\tau, \quad t \in I,$$

where P_t denotes the spectral projection for B_t associated with the spectral component $\{\omega_t\}_{t \in I}$.

As an immediate consequence, we obtain new estimates in the subspace perturbation problem recently considered in [5] and [7].

The paper is organized as follows.

In Section 2 we start with recalling basic facts on orthogonal projections and prove an important technical result (see, Corollary 2.2, The Four Projections Lemma).

In Section 3 we deal with smooth paths of projections. As a key result we relate the norm of the difference of the two endpoints of a smooth path and the corresponding arc length (see, Lemma 3.4, the Arcsine Law for smooth paths).

In Section 4 we provide a characterization of the local geodesic structure of the length space (\mathcal{P}, ρ) and prove that the metric space (\mathcal{P}, ρ) is $\pi/2$ -geodesic. In particular, we generalize the Arcsine Law from Section 3 to the case of continuous paths.

In Section 5 we apply the results from the preceding sections to the problem of variation of spectral subspaces including some discussions about the optimality of the obtained estimates.

In Section 6 we obtain new estimates in the subspace perturbation problem sharpening recent results from [5] and [7].

Acknowledgments. The authors are grateful to Vadim Kostykin for stimulating discussions. A.S. would like to express his gratitude to his scientific advisor Vadim Kostykin for introducing him to the field of the research and continuous support. K.A.M. is indebted to the Institute for Mathematics for its kind hospitality during his stay at the Johannes Gutenberg-Universität Mainz in the Summer of 2009 and 2010. The work of K.A.M. has been supported in part by the Deutsche Forschungsgesellschaft and by the Inneruniversitären Forschungsförderung of the Johannes Gutenberg-Universität Mainz.

2. PRELIMINARIES

We start with recalling some important facts on the representation for the range of an orthogonal projection as a graph subspace associated with the range of another orthogonal projection. For the proofs the reader is referred to the work [6].

Let P and Q be orthogonal projections in the Hilbert space \mathcal{H} , where we will tacitly understand \mathcal{H} to be separable throughout this paper. It is well known that the inequality $\|P - Q\| < 1$ holds true if and only if $\text{Ran } Q$ is a graph of a bounded operator $X \in \mathcal{L}(\text{Ran } P, \text{Ran } P^\perp)$, $P^\perp := I_{\mathcal{H}} - P$, that is,

$$\text{Ran } Q = \mathcal{G}(X) := \mathcal{G}(\text{Ran } P, X) := \{x_0 \oplus Xx_0 \mid x_0 \in \text{Ran } P\}.$$

In this case the projection Q has the following representation as a block operator matrix with respect to the orthogonal decomposition $\mathcal{H} = \text{Ran } P \oplus \text{Ran } P^\perp$:

$$(2.1) \quad Q = \begin{pmatrix} (I_{\mathcal{H}_0} + X^*X)^{-1} & (I_{\mathcal{H}_0} + X^*X)^{-1}X^* \\ X(I_{\mathcal{H}_0} + X^*X)^{-1} & X(I_{\mathcal{H}_0} + X^*X)^{-1}X^* \end{pmatrix},$$

where $\mathcal{H}_0 := \text{Ran } P$ (cf. Remark 3.6 in [6]).

The knowledge of the angular operator X and/or the operator angle Θ (see, e.g., [6] for a discussion of this concept) between the subspaces $\text{Ran } P$ and $\text{Ran } Q$ given by

$$\Theta = \arctan \sqrt{X^*X},$$

provides complete information on relative geometry of the subspaces $\text{Ran } P$ and $\text{Ran } Q$. In particular,

$$(2.2) \quad \|X\| = \frac{\|P - Q\|}{\sqrt{1 - \|P - Q\|^2}} = \tan \|\Theta\|$$

and

$$(2.3) \quad \|P - Q\| = \frac{\|X\|}{\sqrt{1 + \|X\|^2}} = \sin \|\Theta\|$$

(see, e.g., Corollary 3.4 in [6]).

Moreover, in this case, the orthogonal projections P and Q are unitarily equivalent. In particular,

$$P = U^*QU,$$

where U is given by the following unitary block operator matrix

$$(2.4) \quad U = \begin{pmatrix} (I_{\mathcal{H}_0} + X^*X)^{-1/2} & -X^*(I_{\mathcal{H}_1} + XX^*)^{-1/2} \\ X(I_{\mathcal{H}_0} + X^*X)^{-1/2} & (I_{\mathcal{H}_1} + XX^*)^{-1/2} \end{pmatrix}, \quad \mathcal{H}_1 := \text{Ran } P^\perp.$$

Our next result is a purely algebraic observation the proof of which requires nothing but straightforward multiplication of several operator matrices and hence will be omitted.

Lemma 2.1 (Four projections lemma). *Assume that P , Q_1 , and Q_2 are orthogonal projections such that*

$$\|P - Q_j\| < 1, \quad j = 1, 2,$$

and therefore

$$\text{Ran } Q_j = \mathcal{G}(X_j), \quad j = 1, 2,$$

*for some angular operators $X_j \in \mathcal{L}(\text{Ran } P, \text{Ran } P^\perp)$. Let U_1 be the corresponding unitary operator from (2.4) such that $P = U_1^*Q_1U_1$, that is*

$$U_1 = \begin{pmatrix} (I_{\mathcal{H}_0} + X_1^*X_1)^{-1/2} & -X_1^*(I_{\mathcal{H}_1} + X_1X_1^*)^{-1/2} \\ X_1(I_{\mathcal{H}_0} + X_1^*X_1)^{-1/2} & (I_{\mathcal{H}_1} + X_1X_1^*)^{-1/2} \end{pmatrix}$$

with $\mathcal{H}_0 = \text{Ran } P$ and $\mathcal{H}_1 = \text{Ran } P^\perp$. Then the orthogonal projection Q given by

$$Q = U_1^*Q_2U_1,$$

admits the factorization

$$Q = A^{-1/2}BCB^*A^{-1/2},$$

where $A \in \mathcal{L}(\mathcal{H})$, $B \in \mathcal{L}(\text{Ran } P, \text{Ran } P^\perp)$ and $C \in \mathcal{L}(\text{Ran } P)$ are 2×2 , 2×1 and 1×1 block operator matrices (with respect to the orthogonal decomposition $\mathcal{H} = \text{Ran } P \oplus \text{Ran } P^\perp$)

respectively, given by

$$(2.5) \quad A = \begin{pmatrix} I_{\mathcal{H}_0} + X_1^* X_1 & 0 \\ 0 & I_{\mathcal{H}_1} + X_1 X_1^* \end{pmatrix},$$

$$(2.6) \quad B = \begin{pmatrix} I_{\mathcal{H}_0} + X_1^* X_2 \\ X_2 - X_1 \end{pmatrix},$$

$$(2.7) \quad C = (I_{\mathcal{H}_0} + X_2^* X_2)^{-1}.$$

The last statement of this preliminary section allows one to compare the angular operators X_1 and X_2 associated with the graph subspaces $\text{Ran } Q_1$ and $\text{Ran } Q_2$ referred to in Lemma 2.1. As a result, one obtains the following “angle addition” formula.

Corollary 2.2. *Suppose in addition to the assumptions of Lemma 2.1 that the range of the orthogonal projection Q is a graph subspace with respect to the decomposition $\mathcal{H} = \text{Ran } P \oplus \text{Ran } P^\perp =: \mathcal{H}_0 \oplus \mathcal{H}_1$, and therefore*

$$\text{Ran } Q = \mathcal{G}(Z) \quad \text{for some } Z \in \mathcal{L}(\text{Ran } P, \text{Ran } P^\perp).$$

Moreover, assume that the operator $I_{\mathcal{H}_0} + X_2^* X_1 \in \mathcal{L}(\text{Ran } P)$ is of full range, that is,

$$\text{Ran}(I_{\mathcal{H}_0} + X_2^* X_1) = \text{Ran } P.$$

Then

$$(2.8) \quad X_2 - X_1 = (I_{\mathcal{H}_1} + X_1 X_1^*)^{1/2} Z (I_{\mathcal{H}_0} + X_1^* X_1)^{-1/2} (I_{\mathcal{H}_0} + X_1^* X_2).$$

Proof. From the definition of the angular operator Z , i.e. $\text{Ran } Q = \mathcal{G}(Z)$, it follows that

$$(2.9) \quad P^\perp Q = Z P Q.$$

Recall that by Lemma 2.1,

$$Q = A^{-1/2} B C B^* A^{-1/2},$$

where the operators A , B , and C are given by (2.5)-(2.7). In particular,

$$C B^* A^{-1/2}|_{\text{Ran } P} = (I_{\mathcal{H}_0} + X_2^* X_2)^{-1} (I_{\mathcal{H}_0} + X_2^* X_1) (I_{\mathcal{H}_0} + X_1^* X_1)^{-1/2}.$$

By hypothesis, the operator $(I_{\mathcal{H}_0} + X_2^* X_1)$ is of full range, so is $C B^* A^{-1/2}|_{\text{Ran } P}$. Therefore, (2.9) implies the equality

$$P^\perp A^{-1/2} B = Z P A^{-1/2} B.$$

Taking into account representations (2.5) and (2.6), one computes

$$(2.10) \quad P^\perp A^{-1/2} B = (I_{\mathcal{H}_1} + X_1 X_1^*)^{-1/2} (X_2 - X_1)$$

and

$$(2.11) \quad P A^{-1/2} B = (I_{\mathcal{H}_0} + X_1^* X_1)^{-1/2} (I_{\mathcal{H}_0} + X_1^* X_2).$$

Combining (2.9), (2.10), and (2.11), one concludes that

$$(2.12) \quad (I_{\mathcal{H}_1} + X_1 X_1^*)^{-1/2} (X_2 - X_1) = Z (I_{\mathcal{H}_0} + X_1^* X_1)^{-1/2} (I_{\mathcal{H}_0} + X_1^* X_2)$$

and the claim follows by multiplying both sides of (2.12) by the operator $(I_{\mathcal{H}_1} + X_1 X_1^*)^{1/2}$ from the left. \square

Remark 2.3. *Representation (2.8) relating the angular operators X_1 , X_2 and Z is a non-commutative variant of the “angle addition” formula*

$$(2.13) \quad \tan \Theta_2 - \tan \Theta_1 = \tan(\Theta_2 - \Theta_1) \cdot (1 + \tan \Theta_1 \tan \Theta_2).$$

To “justify” this observation, consider an example of a space of dimension 2 and rank 1 orthogonal projections Q_1 and Q_2 whose ranges are lines of inclinations Θ_1 and Θ_2 , respectively. Then the operator Q_2 in the new coordinate system

$$\begin{aligned} x' &= \cos \Theta_1 x + \sin \Theta_1 y \\ y' &= -\sin \Theta_1 x + \cos \Theta_1 y \end{aligned}$$

turns out to be a rank 1 projection Q whose range is a line of the slope $\tan(\Theta_2 - \Theta_1)$. Since the angular operators X_1 , X_2 , and Z play the role of the slope of the line, in the case in question the angle addition formula (2.13) is equivalent to the relation (2.8).

3. SMOOTH PATHS OF PROJECTIONS

Throughout this section we consider the set of orthogonal projections \mathcal{P} in a Hilbert space \mathcal{H} ,

$$\mathcal{P} = \{P \in \mathcal{L}(\mathcal{H}) \mid P = P^* = P^2\},$$

as a metric space with respect to the metric d induced by the operator norm on $\mathcal{L}(\mathcal{H})$.

Recall, that a piecewise C^1 -smooth path is a mapping $\gamma: [a, b] \rightarrow \mathcal{P}$ such that there is a partition $a = t_0 < \dots < t_n = b$ and $\gamma|_{[t_j, t_{j+1}]}$ is C^1 -smooth for all $j \in \{0, \dots, n-1\}$. In particular, such paths are continuous.

Hypothesis 3.1. Assume that $\gamma: [a, b] \rightarrow \mathcal{P}$ is a piecewise C^1 -smooth path of orthogonal projections. Suppose, in addition, that $\gamma(t)$ for all $t \in [a, b]$ is a graph subspace associated with a bounded operator with respect to the orthogonal decomposition $\mathcal{H} = \text{Ran } \gamma(a) \oplus \text{Ran } \gamma(a)^\perp$, that is

$$\text{Ran } \gamma(t) = \mathcal{G}(X_t) \quad \text{for some angular operator } X_t \in \mathcal{L}(\text{Ran } \gamma(a), \text{Ran } \gamma(a)^\perp), \quad t \in [a, b].$$

Our first result in this section shows, that smoothness of the path of projections implies smoothness of the corresponding angular operators in the graph subspace representation. The exact statement is as follows.

Lemma 3.2. Assume Hypothesis 3.1. If $I \subset [a, b]$ is an interval, such that $\gamma|_I$ is C^1 -smooth, then $I \ni t \mapsto X_t$ is also C^1 -smooth. In particular, the path $[a, b] \ni t \mapsto X_t$ is piecewise C^1 -smooth.

Proof. Let $\mathcal{H}_0 = \text{Ran } \gamma(a)$ and $\mathcal{H}_1 = \text{Ran } \gamma(a)^\perp$, where $\gamma(a)^\perp = I_{\mathcal{H}} - \gamma(a)$. Introduce piecewise C^1 -smooth families of bounded operators given by

$$T_t := \gamma(a)^\perp \gamma(t) \gamma(a) \quad \text{and} \quad S_t := \gamma(a) \gamma(t) \gamma(a), \quad t \in [a, b].$$

Denote by R_t the following operator matrix with respect to the decomposition $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$,

$$(3.1) \quad R_t := \begin{pmatrix} I_{\mathcal{H}_0} + X_t^* X_t & 0 \\ 0 & 0 \end{pmatrix}, \quad t \in [a, b].$$

Using (2.1), one obtains that for each $t \in [a, b]$

$$(3.2) \quad \gamma(t) = \begin{pmatrix} (I_{\mathcal{H}_0} + X_t^* X_t)^{-1} & (I_{\mathcal{H}_0} + X_t^* X_t)^{-1} X_t^* \\ X_t (I_{\mathcal{H}_0} + X_t^* X_t)^{-1} & X_t (I_{\mathcal{H}_0} + X_t^* X_t)^{-1} X_t^* \end{pmatrix},$$

and a simple computation shows that the operators T_t and S_t can be represented as the following operator matrices with respect to the decomposition $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$:

$$\begin{aligned} T_t &= \gamma(a)^\perp \gamma(t) \gamma(a) = \begin{pmatrix} 0 & 0 \\ X_t (I_{\mathcal{H}_0} + X_t^* X_t)^{-1} & 0 \end{pmatrix}, \\ S_t &= \gamma(a) \gamma(t) \gamma(a) = \begin{pmatrix} (I_{\mathcal{H}_0} + X_t^* X_t)^{-1} & 0 \\ 0 & 0 \end{pmatrix}, \quad t \in [a, b] \end{aligned}$$

Now it is easy to see that the “resolvent identity”

$$(3.3) \quad R_s - R_t = R_t (S_t - S_s) R_s, \quad s, t \in [a, b].$$

holds.

The norm estimate

$$\|R_t\| \leq 1 + \|X_t\|^2$$

combined with the identity (see (2.2))

$$\|X_t\| = \frac{\|\gamma(t) - \gamma(a)\|}{\sqrt{1 - \|\gamma(t) - \gamma(a)\|^2}}$$

yields the inequality

$$(3.4) \quad \|R_t\| \leq \frac{1}{1 - \|\gamma(t) - \gamma(a)\|^2}, \quad t \in [a, b].$$

Since by hypothesis $\text{Ran } \gamma(t)$ is a graph of a bounded operator, one gets that $\|\gamma(t) - \gamma(a)\| < 1$ for all $t \in [a, b]$. Due to the continuity of the path $[a, b] \ni t \mapsto \gamma(t)$, from (3.4) one concludes that for any $t_0 \in [a, b]$ there exists a neighborhood U_{t_0} of the point t_0 such that the function $U_{t_0} \ni t \mapsto \|R_t\|$ is uniformly bounded. Taking this observation into account and recalling that the family S_t is piecewise differentiable, from the representation (3.3) it follows that the family R_t is also piecewise differentiable with

$$(3.5) \quad \dot{R}_t = -R_t \dot{S}_t R_t, \quad t \in I,$$

where $I \subset [a, b]$ is any interval such that $\gamma|_I$ is C^1 -smooth. Since $I \ni t \mapsto \dot{S}_t$ is a continuous path, from (3.5) it follows that $I \ni t \mapsto R_t$ is a C^1 -smooth path. It remains to observe that

$$\begin{pmatrix} 0 & 0 \\ X_t & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ X_s & 0 \end{pmatrix} = T_t R_t - T_s R_s, \quad s, t \in I,$$

to conclude that $I \ni t \mapsto X_t$ is a C^1 -smooth path with

$$\begin{pmatrix} 0 & 0 \\ \dot{X}_t & 0 \end{pmatrix} = \dot{T}_t R_t + T_t \dot{R}_t, \quad t \in I.$$

□

Our next result forms in fact the core of our considerations for it relates the evolution of the path of angular operators and the evolution of the corresponding path of orthogonal projections. It justifies the following principle: The speed of rotation of the subspaces $\text{Ran } \gamma(t)$ along a path $[a, b] \ni t \mapsto \gamma(t)$ does not exceed the speed on the path.

Lemma 3.3. *Assume Hypothesis 3.1. Let $I \subset [a, b]$ be an interval, such that $\gamma|_I$ is C^1 -smooth. Then the estimate*

$$(3.6) \quad \|\dot{X}_t\| \leq (1 + \|X_t\|^2) \|\dot{\gamma}(t)\|$$

holds for all $t \in I$.

Proof. Since, by hypothesis, $\|\gamma(t) - \gamma(a)\| < 1$, $t \in I$, the projections $\gamma(t)$ and $\gamma(a)$ are unitarily equivalent. In particular,

$$(3.7) \quad \gamma(a) = U_t^* \gamma(t) U_t, \quad t \in I.$$

where the family of unitary operators U_t , $t \in I$, is given by (2.4) accordingly.

Fix an $s \in I$ and introduce the family of orthogonal projections

$$(3.8) \quad Q_t = U_t^* \gamma(s) U_t, \quad t \in I.$$

Due to the continuity of the path $I \ni t \mapsto \gamma(t)$, there exists a neighborhood $\mathcal{V} \subset I$ of the point s , such that

$$(3.9) \quad \|\gamma(t) - \gamma(s)\| < 1, \quad t \in \mathcal{V}.$$

Since by (3.7) and (3.8)

$$(3.10) \quad \|Q_t - \gamma(a)\| = \|U_t^* \gamma(s) U_t - U_t^* \gamma(t) U_t\| = \|\gamma(s) - \gamma(t)\|,$$

from (3.9) it follows that

$$\|Q_t - \gamma(a)\| < 1, \quad t \in \mathcal{V}.$$

Therefore, $\text{Ran } Q_t$ is a graph subspace with respect to the decomposition $\mathcal{H} = \text{Ran } \gamma(a) \oplus \text{Ran } \gamma(a)^\perp$, that is

$$\text{Ran } Q_t = \mathcal{G}(Y_t) \quad \text{for some } Y_t \in \mathcal{L}(\text{Ran } \gamma(a), \text{Ran } \gamma(a)^\perp).$$

Next, one observes that the operator $I_{\mathcal{H}_0} + X_s^* X_s$ has a bounded inverse and therefore $I_{\mathcal{H}_0} + X_t^* X_s$ has a bounded inverse as well for all t from in a possibly smaller neighborhood $\tilde{\mathcal{V}} \subset \mathcal{V}$ of the point s . In particular, the operator $I_{\mathcal{H}_0} + X_t^* X_s$ is of full range for all $t \in \tilde{\mathcal{V}}$ and one can apply Corollary 2.2 to get the representation

$$X_t - X_s = (I_{\mathcal{H}_0} + X_s X_s^*)^{1/2} Y_t (I_{\mathcal{H}_0} + X_s^* X_s)^{-1/2} (I_{\mathcal{H}_0} + X_s^* X_t), \quad t \in \tilde{\mathcal{V}},$$

and hence

$$(3.11) \quad \|X_t - X_s\| \leq \|(I + X_s X_s^*)^{1/2}\| \cdot \|Y_t\| \cdot \|(I + X_s^* X_s)^{-1/2} (I + X_s^* X_t)\|, \quad t \in \tilde{\mathcal{V}}.$$

Since by (2.2)

$$\|Y_t\| = \frac{\|Q_t - \gamma(a)\|}{\sqrt{1 - \|Q_t - \gamma(a)\|^2}}$$

from (3.10) one obtains that

$$\lim_{t \rightarrow s} \frac{\|Y_t\|}{t - s} = \lim_{t \rightarrow s} \frac{\|\gamma(t) - \gamma(s)\|}{t - s} = \|\dot{\gamma}(s)\|,$$

for $I \ni t \mapsto \gamma(t)$ is a C^1 -smooth path. Since by Lemma 3.2 $I \ni t \mapsto X_t$ is also a C^1 -smooth path, from inequality (3.11) one gets the estimate

$$\begin{aligned} \|\dot{X}_s\| &\leq \|(I + X_s X_s^*)^{1/2}\| \cdot \|\dot{\gamma}(s)\| \cdot \|(I + X_s^* X_s)^{1/2}\| \\ &= (1 + \|X_s\|^2) \cdot \|\dot{\gamma}(s)\|. \end{aligned}$$

Since the reference point $s \in I$ has been chosen arbitrarily, one proves the inequality (3.6). \square

Using the information about the evolution of the angular operators provided by Lemma 3.3, we are now able to estimate the variation of the corresponding orthogonal projections.

Lemma 3.4 (The Arcsine Law). *Let $\gamma: [a, b] \rightarrow \mathcal{P}$ be a piecewise C^1 -smooth path. Then*

$$(3.12) \quad \arcsin(\|\gamma(b) - \gamma(a)\|) \leq l_R(\gamma),$$

where

$$l_R(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt$$

is the Riemannian length of the path γ .

Proof. Since $\|\gamma(b) - \gamma(a)\| \leq 1$, we may assume $l_R(\gamma) < \frac{\pi}{2}$.

Let $a = t_0 < \dots < t_n = b$ be a partition such that $\gamma|_{[t_j, t_{j+1}]}$ is C^1 -smooth. Set

$$(3.13) \quad T := \sup \{t \in [a, b] \mid \|\gamma(t') - \gamma(a)\| < 1 \text{ for all } t' \in [a, t]\}.$$

Clearly, $T > a$, for γ is continuous. Since

$$\|\gamma(t) - \gamma(a)\| < 1 \quad \text{for all } t \in [a, T],$$

the range of $\gamma(t)$ is a graph subspace with respect to the decomposition $\mathcal{H} = \text{Ran } \gamma(a) \oplus \text{Ran } \gamma(a)^\perp$ and therefore

$$\text{Ran } \gamma(t) = \mathcal{G}(X_t) \quad \text{for some } X_t \in \mathcal{L}(\text{Ran } \gamma(a), \text{Ran } \gamma(a)^\perp), \quad t \in [a, T].$$

Due to Lemma 3.3, we have

$$\|\dot{X}_t\| \leq (1 + \|X_t\|^2) \|\dot{\gamma}(t)\|$$

for all $t \in (t_j, t_{j+1})$, $j = 0, \dots, n-1$, as long as $t < T$. For arbitrary $t \in [a, T]$ there is a unique $k \in \{0, \dots, n-1\}$ such that $t \in [t_k, t_{k+1})$. We obtain

$$\begin{aligned} \|X_t\| &= \|X_t - X_a\| \leq \|X_t - X_{t_k}\| + \sum_{j=0}^{k-1} \|X_{t_{j+1}} - X_{t_j}\| \\ &\leq \int_{t_k}^t \|\dot{X}_\tau\| d\tau + \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \|\dot{X}_\tau\| d\tau \\ (3.14) \quad &\leq \int_{t_k}^t (1 + \|X_\tau\|^2) \|\dot{\gamma}(\tau)\| d\tau + \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} (1 + \|X_\tau\|^2) \|\dot{\gamma}(\tau)\| d\tau \\ &= \int_a^t (1 + \|X_\tau\|^2) \|\dot{\gamma}(\tau)\| d\tau \end{aligned}$$

for $t \in [a, T]$. Denoting the right hand side of (3.14) by $F(t)$, i.e.

$$(3.15) \quad F(t) = \int_0^t (1 + \|X_\tau\|^2) \|\dot{\gamma}(\tau)\| d\tau,$$

one concludes that

$$(3.16) \quad \|X_t\| \leq F(t), \quad t \in [a, T],$$

and hence

$$(3.17) \quad F'(t) = (1 + \|X_t\|^2) \|\dot{\gamma}(t)\| \leq (1 + F^2(t)) \|\dot{\gamma}(t)\|$$

for $t \in [a, T)$ except for the finitely many points t_j . Since by assumption $l_R(\gamma) < \frac{\pi}{2}$, one can solve the differential inequality (3.17) on every sub-interval of $[a, T)$ where F is C^1 -smooth. For $t \in [t_k, t_{k+1}]$, $t < T$, one obtains

$$\begin{aligned} \arctan F(t) &= \arctan F(t) - \arctan F(a) \\ &= \arctan F(t) - \arctan F(t_k) + \sum_{j=0}^{k-1} (\arctan F(t_{j+1}) - \arctan F(t_j)) \\ &\leq \int_{t_k}^t \|\dot{\gamma}(\tau)\| d\tau + \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \|\dot{\gamma}(\tau)\| d\tau = \int_a^t \|\dot{\gamma}(\tau)\| d\tau. \end{aligned}$$

Together with (3.16) this yields the bound

$$(3.18) \quad \arctan \|X_t\| \leq \int_a^t \|\dot{\gamma}(\tau)\| d\tau \leq l_R(\gamma) < \frac{\pi}{2}, \quad t \in [a, T].$$

Since

$$\arcsin(\|\gamma(t) - \gamma(a)\|) = \arctan(\|X_t\|), \quad t \in [a, T],$$

from (3.18) one gets the estimate

$$\arcsin(\|\gamma(t) - \gamma(a)\|) \leq \int_a^t \|\dot{\gamma}(\tau)\| d\tau, \quad t \in [a, T],$$

and hence, by continuity,

$$\arcsin(\|\gamma(t) - \gamma(a)\|) \leq \int_0^t \|\dot{\gamma}(\tau)\| d\tau \leq l_R(\gamma) < \frac{\pi}{2}, \quad t \in [a, T].$$

In particular it is $T = b$ by definition of T in (3.13), which proves (3.12). \square

The next lemma shows that the inequality of Lemma 3.4 is sharp. In particular, it states that given orthogonal projections P and Q with $\|P - Q\| < 1$, one can construct a C^1 -smooth path of minimal length, a geodesic, among all (C^1 -smooth) paths connecting P and Q . It will turn out later, that this same path is of minimal length even among all continuous paths connecting P and Q .

Lemma 3.5. *Let $P, Q \in \mathcal{P}$ with $\|P - Q\| < 1$. Then there exists a C^1 -smooth path $\gamma: [0, l] \rightarrow \mathcal{P}$ connecting P and Q such that*

$$\|\dot{\gamma}(t)\| = 1, \quad t \in [0, l],$$

where

$$l = \arcsin(\|P - Q\|) < \frac{\pi}{2}.$$

Proof. Since $\|P - Q\| < 1$, the range of Q is a graph subspace with respect to $\text{Ran } P$, i.e. $\text{Ran } Q = \mathcal{G}(\text{Ran } P, X)$ for some $X \in \mathcal{L}(\text{Ran } P, \text{Ran } P^\perp)$. Without loss of generality one can assume that the pair (P, Q) is generic, that is,

$$\text{Ran } P \cap \text{Ran } Q = \text{Ran } P^\perp \cap \text{Ran } Q^\perp = \{0\}$$

and hence one can write (see [7, Theorem 2.2])

$$P = \begin{pmatrix} I_{\text{Ran } P} & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = \mathcal{W}^* \begin{pmatrix} \cos^2 \Theta & \sin \Theta \cos \Theta \\ \sin \Theta \cos \Theta & \sin^2 \Theta \end{pmatrix} \mathcal{W}$$

with respect to the decomposition $\mathcal{H} = \text{Ran } P \oplus \text{Ran } P^\perp$, where $\Theta = \arctan \sqrt{X^* X}$ is the corresponding operator angle and \mathcal{W} is a unitary operator. In particular,

$$l := \|\Theta\| = \arcsin(\|P - Q\|) < \frac{\pi}{2}.$$

Introduce the C^1 -smooth path $\gamma: [0, l] \rightarrow \mathcal{P}$ connecting P to Q by the following family of block operator matrices with respect to the decomposition $\mathcal{H} = \text{Ran } P \oplus \text{Ran } P^\perp$:

$$\gamma(t) = \mathcal{W}^* \begin{pmatrix} \cos^2(\Theta_t^t) & \sin(\Theta_t^t) \cos(\Theta_t^t) \\ \sin(\Theta_t^t) \cos(\Theta_t^t) & \sin^2(\Theta_t^t) \end{pmatrix} \mathcal{W}, \quad t \in [0, l].$$

It remains to observe that

$$\begin{aligned} \dot{\gamma}(t) &= \frac{1}{l} \mathcal{W}^* \begin{pmatrix} -\Theta \sin(2\Theta_t^t) & \Theta \cos(2\Theta_t^t) \\ \Theta \cos(2\Theta_t^t) & \Theta \sin(2\Theta_t^t) \end{pmatrix} \mathcal{W} \\ &= \frac{1}{l} \mathcal{W}^* \begin{pmatrix} \Theta^{1/2} & 0 \\ 0 & \Theta^{1/2} \end{pmatrix} J \begin{pmatrix} \Theta^{1/2} & 0 \\ 0 & \Theta^{1/2} \end{pmatrix} \mathcal{W}, \quad t \in [0, l], \end{aligned}$$

where J is a self-adjoint involution, $J^2 = I$, given by

$$J = \begin{pmatrix} -\sin(2\Theta_t^t) & \cos(2\Theta_t^t) \\ \cos(2\Theta_t^t) & \sin(2\Theta_t^t) \end{pmatrix}, \quad t \in [0, l].$$

Therefore,

$$\|\dot{\gamma}(t)\| = \frac{\|\Theta\|}{l} = 1, \quad t \in [0, l],$$

which completes the proof. \square

4. THE SPACE \mathcal{P} AS A LOCAL GEODESIC METRIC SPACE

The main goal of this section is to study metric properties of the space \mathcal{P} considered as a length space.

Recall necessary definitions. Given a continuous path $\gamma: [a, b] \rightarrow \mathcal{P}$, its length $l(\gamma)$ is defined by

$$l(\gamma) = \sup \left\{ \sum_{j=0}^{n-1} \|\gamma(t_{j+1}) - \gamma(t_j)\| \mid n \in \mathbb{N}, a = t_0 < \dots < t_n = b \right\}.$$

Recall that a path is called rectifiable if its length is finite.

On \mathcal{P} introduce a length or inner (pseudo-)metric ρ given by the formula

$$\rho(P, Q) = \text{infimum of length of rectifiable paths } \gamma \text{ from } P \text{ to } Q.$$

If there are no such paths then set $\rho(P, Q) = \infty$.

It is well known that the pseudometric ρ is actually a metric and therefore (\mathcal{P}, ρ) is a well defined metric space (cf., [2, Proposition 3.2], the length space).

There is another way of introducing the inner metric via a “Riemannian” arc length of piecewise differentiable paths $\gamma: [a, b] \rightarrow \mathcal{P}$ from P to Q by

$$\rho_R(P, Q) = \text{infimum of } \int_a^b \|\dot{\gamma}(t)\| dt,$$

and $\rho_R(P, Q) = \infty$ if there are no such paths.

The following lemma shows that the metric spaces (\mathcal{P}, ρ) and (\mathcal{P}, ρ_R) coincide.

Lemma 4.1. *Let $\gamma: [a, b] \rightarrow \mathcal{P}$ be a continuous path. Then there is a sequence (γ_n) of piecewise C^1 -smooth paths $\gamma_n: [a, b] \rightarrow \mathcal{P}$, each having the same endpoints as γ , such that γ_n converges uniformly to γ and its length $l(\gamma_n)$ converges to $l(\gamma)$. In particular, the inner metric ρ and the Riemannian pseudometric ρ_R on \mathcal{P} coincide.*

Proof. Since $\gamma: [a, b] \rightarrow \mathcal{P}$ is uniformly continuous, we can choose for each $n \in \mathbb{N}$ some $N(n) \in \mathbb{N}$ and a partition $a = t_0^{(n)} < \dots < t_{N(n)}^{(n)} = b$ such that

$$(4.1) \quad \|\gamma(t) - \gamma(s)\| < \frac{1}{n}$$

for all $t, s \in [t_j^{(n)}, t_{j+1}^{(n)}]$, $j \in \{0, \dots, N(n) - 1\}$. By Lemma 3.5 we can choose C^1 -smooth paths in \mathcal{P} connecting $\gamma(t_j^{(n)})$ and $\gamma(t_{j+1}^{(n)})$ with length $\arcsin(\|\gamma(t_{j+1}^{(n)}) - \gamma(t_j^{(n)})\|)$ for $j \in \{0, \dots, N(n)\}$. Let $\gamma_n: [a, b] \rightarrow \mathcal{P}$ denote the concatenation of these paths for every $n \in \mathbb{N}$. Obviously, each γ_n is piecewise C^1 -smooth and has the same endpoints as γ . Since $\gamma_n(t_j^{(n)}) = \gamma(t_j^{(n)})$ we have in addition for every $t \in [t_j^{(n)}, t_{j+1}^{(n)}]$ the estimate

$$\begin{aligned} \|\gamma_n(t) - \gamma(t)\| &\leq \|\gamma_n(t) - \gamma_n(t_j^{(n)})\| + \|\gamma(t_j^{(n)}) - \gamma(t)\| \\ &\leq l(\gamma_n|_{[t_j^{(n)}, t_{j+1}^{(n)}]}) + \|\gamma(t_j^{(n)}) - \gamma(t)\| \\ &< \arcsin\left(\frac{1}{n}\right) + \frac{1}{n}, \end{aligned}$$

and therefore

$$\|\gamma_n(t) - \gamma(t)\| < \arcsin\left(\frac{1}{n}\right) + \frac{1}{n}$$

for all $t \in [a, b]$, i.e. γ_n converges uniformly to γ .

In order to show that $l(\gamma_n)$ converges to $l(\gamma)$, let $\varepsilon > 0$ be arbitrary and take $k \in \mathbb{N}$ such that $\frac{l(\gamma)}{k} < \varepsilon$. Since $\frac{\arcsin(x)}{x}$ goes to 1 as x approaches zero, there is some $\delta > 0$ such that

$$(4.2) \quad \arcsin(x) \leq \left(1 + \frac{1}{k}\right) x$$

for all $0 \leq x < \delta$. Due to the lower semicontinuity of the length of paths (cf., [2, Proposition 1.20]) we can take $N \in \mathbb{N}$ such that

$$(4.3) \quad l(\gamma) \leq l(\gamma_n) + \varepsilon$$

whenever $n \geq N$. We may assume that $\frac{1}{N} < \delta$. Taking (4.1) into account, from (4.2) and the additivity of the length of paths one obtains

$$\begin{aligned} l(\gamma_n) &= \sum_{j=0}^{N(n)-1} \arcsin(\|\gamma(t_{j+1}^{(n)}) - \gamma(t_j^{(n)})\|) \leq \left(1 + \frac{1}{k}\right) \sum_{j=0}^{N(n)-1} \|\gamma(t_{j+1}^{(n)}) - \gamma(t_j^{(n)})\| \\ &\leq \left(1 + \frac{1}{k}\right) \cdot l(\gamma) \end{aligned}$$

and therefore

$$l(\gamma_n) - l(\gamma) \leq \frac{l(\gamma)}{k} < \varepsilon.$$

for all $n \geq N$. Together with (4.3) we arrive at

$$|l(\gamma_n) - l(\gamma)| \leq \varepsilon$$

whenever $n \geq N$, i.e. $l(\gamma_n)$ converges to $l(\gamma)$, which completes the proof. \square

As a consequence, we may restrict our further considerations to piecewise C^1 -smooth paths only. The continuous case follows from that by approximation with piecewise smooth paths. In particular, we can relax the smoothness hypothesis of Lemma 3.4 and obtain the following result, the Arcsine Law for continuous paths.

Corollary 4.2. *Let $\gamma: [a, b] \rightarrow \mathcal{P}$ be a continuous path. Then*

$$(4.4) \quad \arcsin(\|\gamma(b) - \gamma(a)\|) \leq l(\gamma).$$

Recall that given a metric space (X, d) , a geodesic path joining x to y is a map γ from a closed interval $[0, l]$ to X such that $\gamma(0) = x$, $\gamma(l) = y$ and $\rho(\gamma(t), \gamma(s)) = |s - t|$ for all $s, t \in [0, l]$. We also recall that a metric space (X, d) is said to be r -geodesic if for every pair of points $x, y \in X$ with $d(x, y) < r$ there is a geodesic path joining x to y .

The main result of this geometric section characterizes the local geodesic behavior of the length space (\mathcal{P}, ρ) . In particular, we obtain a concrete local representation of the length metric ρ in terms of the norm of the angle operator.

Theorem 4.3. *The metric space (\mathcal{P}, ρ) is $\frac{\pi}{2}$ -geodesic. Moreover, it is $\rho(P, Q) < \frac{\pi}{2}$ if and only if $d(P, Q) = \|P - Q\| < 1$. In that case*

$$\rho(P, Q) = \arcsin(\|P - Q\|).$$

In particular, $\text{Ran } Q$ is a graph subspace with respect to the orthogonal decomposition $\mathcal{H} = \text{Ran } P \oplus \text{Ran } P^\perp$ and

$$\rho(P, Q) = \|\Theta\|,$$

where Θ is the operator angle between $\text{Ran } Q$ and $\text{Ran } P$.

Proof. Suppose that P and Q are orthogonal projections such that $\rho(P, Q) < \frac{\pi}{2}$. In particular, since $\rho(P, Q)$ is finite, this means that there is a continuous path γ connecting P and Q . For any such path we have by Corollary 4.2 that

$$(4.5) \quad \arcsin(\|P - Q\|) \leq l(\gamma).$$

Going to the infimum over connecting paths, we obtain

$$(4.6) \quad \arcsin(\|P - Q\|) \leq \rho(P, Q),$$

and hence

$$\|P - Q\| \leq \sin(\rho(P, Q)) < 1,$$

due to $\rho(P, Q) < \frac{\pi}{2}$.

Conversely, if $\|P - Q\| < 1$, by Lemma 3.5 there is a C^1 -smooth geodesic path γ connecting P and Q of length $l(\gamma) = \arcsin(\|P - Q\|)$ and therefore

$$(4.7) \quad \rho(P, Q) \leq l(\gamma) = \arcsin(\|P - Q\|) < \frac{\pi}{2}.$$

Thus, (\mathcal{P}, ρ) is $\frac{\pi}{2}$ -geodesic, and combining (4.6) and (4.7) proves the remaining statement of the theorem. \square

5. APPLICATIONS

Paths of orthogonal projections naturally arise when considering families of self-adjoint operators depending smoothly on a parameter. Under the additional hypothesis that the self-adjoint family has a spectrum consisting of two separated parts, the main problem is to obtain integral estimates in terms of the relative strength of the perturbation along the path versus the distance between the components. The upper semicontinuity of the spectrum under a perturbation allows one to obtain efficient estimates on the rotation angle of the spectral subspaces, especially in the case where the *a posteriori* knowledge of the evolution of the separated parts of the spectra is known.

First, we recall the concept of an upper semicontinuous family of sets depending on a parameter.

Definition 5.1. *We say that a family of sets $\{\omega_t\}_{t \in I}$, with I an interval, is upper semicontinuous at the point $t \in I$ if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$(5.1) \quad \rho(\omega_s, \omega_t) = \sup_{\lambda \in \omega_s} \text{dist}(\lambda, \omega_t) < \varepsilon \quad \text{whenever} \quad |s - t| < \delta, \quad s, t \in I.$$

The family $\{\omega_t\}_{t \in I}$, is called upper semicontinuous on I if it is upper semicontinuous at any point $t \in I$.

Without any loss of generality, we will assume any interval I to contain 0 throughout this section.

It is well known (see, e.g., [4, Theorem V.4.10]), that given a C^1 -smooth path $I \ni t \mapsto B_t$ of self-adjoint bounded operators, the family of their spectra $\{\text{spec}(B_t)\}_{t \in I}$ is upper semicontinuous on I . Under the additional assumption that the spectrum of each B_t is separated into two disjoint components, one can expect the two corresponding families of spectral components to be upper semicontinuous as well, provided that they are chosen appropriately. Under these hypotheses, one can study the variation of the corresponding spectral subspaces under a variation of the parameter $t \in I$. A natural way of doing that, is to estimate the deviation of the corresponding spectral projections in the length space (\mathcal{P}, ρ) .

As the main application of Theorem 3.1 we obtain the following result.

Theorem 5.2. Assume that $I \ni t \mapsto B_t$ is a C^1 -smooth path of self-adjoint bounded operators. Suppose that the spectrum of each B_t consists of two disjoint spectral components that upper semicontinuously depend on the parameter t . That is, assume that there exist nonempty closed subsets $\omega_t, \Omega_t \subset \mathbb{R}$ such that for all $t \in I$

- (i) $\text{spec}(B_t) = \omega_t \cup \Omega_t$,
- (ii) $\text{dist}(\omega_t, \Omega_t) > 0$,
- (iii) the families $\{\omega_t\}_{t \in I}$ and $\{\Omega_t\}_{t \in I}$ are upper semicontinuous on I .

Let

$$(5.2) \quad P_t := E_{B_t}(\omega_t), \quad t \in I,$$

denote the spectral projection of the self-adjoint operator B_t associated with the set ω_t .

Then

$$(5.3) \quad \rho(P_t, P_0) = \arcsin(\|P_t - P_0\|) \leq \frac{\pi}{2} \int_0^t \frac{\|\dot{B}_\tau\|}{\text{dist}(\omega_\tau, \Omega_\tau)} d\tau, \quad t \in I.$$

Proof. We make use of the concept of double operator integrals. Those readers, who prefer to see a “standard” proof are referred to Appendix C.

Recall that by the Daletskii-Krein differentiation formula one obtains the representation

$$(5.4) \quad \frac{d}{dt} f(B_t) = \int \int \frac{f(\lambda) - f(\mu)}{\lambda - \mu} dE_{B_t}(\lambda) \dot{B}_t dE_{B_t}(\mu),$$

where dE_{B_t} stands for the spectral measure of the self-adjoint operator B_t and f is a C^∞ -function on an open interval (a, b) containing the spectrum of the bounded operator B_t . Under the spectra separation hypothesis one can find an $f \in C_0^\infty([a, b])$ such that

$$f(\lambda) = \begin{cases} 1, & \lambda \in \omega_t, \\ 0, & \lambda \in \Omega_t. \end{cases}$$

For those f 's one easily concludes that

$$f(B_t) = E_{B_t}(\omega_t) = P_t, \quad t \in I,$$

and therefore, from (5.4), one obtains the representation

$$\dot{P}_t = \int \int \frac{f(\lambda) - f(\mu)}{\lambda - \mu} dE_{B_t}(\lambda) \dot{B}_t dE_{B_t}(\mu), \quad t \in I.$$

Hence,

$$(5.5) \quad \begin{aligned} P_t \dot{P}_t P_t^\perp &= \int \int \frac{f(\lambda) - f(\mu)}{\lambda - \mu} (P_t dE_{B_t}(\lambda)) \dot{B}_t (dE_{B_t}(\mu) P_t^\perp) \\ &= \int \int \frac{1}{\lambda - \mu} (P_t dE_{B_t}(\lambda)) P_t \dot{B}_t P_t^\perp (dE_{B_t}(\mu) P_t^\perp), \quad t \in I. \end{aligned}$$

Since the spectral measures $P_t dE_{B_t}$ and $dE_{B_t}(\mu) P_t^\perp$ are supported by the sets ω_t and Ω_t , respectively, and the sets ω_t and Ω_t are separated with

$$\text{dist}(\omega_t, \Omega_t) > 0,$$

the right hand side of (5.5) can be represented as follows

$$(5.6) \quad \int \int \frac{1}{\lambda - \mu} (P_t dE_{B_t}(\lambda)) P_t \dot{B}_t P_t^\perp (dE_{B_t}(\mu) P_t^\perp) = \int_{\mathbb{R}} e^{isB_t} P_t \dot{B}_t P_t^\perp e^{-isB_t} g(s) ds,$$

where g denotes any function in $L^1(\mathbb{R})$, continuous except at zero, such that

$$\int_{\mathbb{R}} e^{-is\lambda} g(s) ds = \frac{1}{\lambda} \quad \text{whenever} \quad |\lambda| \geq \frac{1}{\text{dist}(\omega_t, \Omega_t)}.$$

In particular, one gets the estimate

$$\|\dot{P}_t\| = \|P_t \dot{P}_t P_t^\perp\| \leq c \frac{\|\dot{P}_t B_t P_t^\perp\|}{\text{dist}(\omega_t, \Omega_t)} \leq c \frac{\|\dot{B}_t\|}{\text{dist}(\omega_t, \Omega_t)}, \quad t \in I,$$

where

$$c = \inf \left\{ \|g\|_{L^1(\mathbb{R})} : g \in L^1(\mathbb{R}), \widehat{g}(\lambda) = \frac{1}{\lambda}, |\lambda| \geq 1 \right\},$$

In fact, see [9],

$$c = \frac{\pi}{2},$$

and hence one gets the estimate

$$(5.7) \quad \|\dot{P}_t\| \leq \frac{\pi}{2} \frac{\|\dot{B}_t\|}{\text{dist}(\omega_t, \Omega_t)}, \quad t \in I.$$

Applying Lemma 3.4 completes the proof. \square

Remark 5.3. We refer to the work of R. McEachin [8] where in fact it is shown that the norm of the transformer given by the double operator integral (5.6) is $\frac{\pi}{2\text{dist}(\omega_t, \Omega_t)}$ and therefore one cannot expect to get an estimate better than (5.7) in general.

However, if, in addition to the hypotheses of Theorem 5.2, the spectral components ω_t and Ω_t are subordinated, i.e. $\sup \omega_t > \inf \Omega_t$, or vice versa, or if they are annular separated, that is, the convex hull of ω_t lies in the complement to the set Ω_t for all $t \in I$, or vice versa, the estimate (5.3) can be strengthened as follows

$$(5.8) \quad \arcsin(\|P_t - P_0\|) \leq \int_0^t \frac{\|\dot{B}_\tau\|}{\text{dist}(\omega_\tau, \Omega_\tau)} d\tau, \quad t \in I.$$

Note, that this estimate is sharp in general (at least in the case of subordinated spectra), as we already know from our previous considerations in sections 3 and 4. Indeed, for a C^1 -smooth path $I \ni t \mapsto P_t$ of orthogonal projections take $B_t = P_t$, $\omega_t = \{1\}$ and $\Omega_t = \{0\}$, $t \in I$. Then it is $\text{spec}(B_t) = \omega_t \cup \Omega_t$ and $\text{dist}(\omega_t, \Omega_t) = 1$ for all $t \in I$ and therefore, in this case, (5.8) coincides with (3.12), which is sharp in general.

The following proposition based on a detailed analysis of one of the realizations of the Heisenberg commutation relations shows that the estimate (5.3) in Theorem 5.2, being understood in a somewhat more general context where the consideration of unbounded operators is not excluded, is sharp.

Proposition 5.4. Let D be the differentiation operator with periodic boundary conditions in $L^2(-1, 1)$ given by the differential expression

$$(5.9) \quad D = \frac{d}{dx} \quad \text{on} \quad \text{Dom}(D) = \{f \in W^{2,1}(-1, 1) \mid f(-1) = f(1)\}.$$

Introduce the isospectral path $[0, \frac{\pi}{2}] \ni t \mapsto B_t$ of unbounded self-adjoint operators

$$(5.10) \quad B_t = U_t(iD)U_t^* \quad \text{on} \quad \text{Dom}(B_t) = U_t \text{Dom}(D),$$

where U_t is the family of unitary operators given by

$$(5.11) \quad (U_t f)(x) = e^{i2tx} f(x), \quad t \in \left[0, \frac{\pi}{2}\right].$$

Set

$$(5.12) \quad \omega_t = 2\pi\mathbb{Z}, \quad \Omega_t = 2\pi\mathbb{Z} \setminus \pi\mathbb{Z},$$

and denote by P_t the spectral projection of B_t onto the subspace of “even harmonics”, that is,

$$(5.13) \quad P_t = E_{B_t}(\omega_t), \quad t \in \left[0, \frac{\pi}{2}\right].$$

Then

$$(5.14) \quad \arcsin(\|P_t - P_0\|) = \frac{\pi}{2} \int_0^t \frac{\|\overline{\dot{B}_\tau}\|}{\text{dist}(\omega_\tau, \Omega_\tau)} d\tau, \quad t \in \left[0, \frac{\pi}{2}\right],$$

where $\overline{\dot{B}_t}$ denotes the closure of the strong derivative $\dot{B}_t = \frac{d}{dt}B_t$ of the path initially defined on

$$\text{Dom}(\dot{B}_t) = C_0^\infty(-1, 1).$$

Proof. First, one observes that

$$(5.15) \quad \dot{B}_t f = \frac{d}{dt}B_t f = 2U_t[\hat{x}, D]U_t^* f = -2f \quad \text{for all } f \in C_0^\infty(-1, 1),$$

where we used the commutation relation

$$[\hat{x}, D] = I \quad \text{on} \quad C_0^\infty(-1, 1) \subset \text{Dom}(\hat{x}D) \cap \text{Dom}(D\hat{x})$$

relating the differentiation operator D and the (bounded) multiplication operator \hat{x} by the independent variable on $L^2(-1, 1)$. Thus, the strong derivative \dot{B}_t is well defined on $\text{Dom}(\dot{B}_t) = C_0^\infty(-1, 1)$ and hence

$$\overline{\dot{B}_t} = -2I_{L^2(-1, 1)}.$$

On the other hand, the spectrum of iD consists of simple eigenvalues located at the points of the lattice $2\pi\mathbb{Z}$, so does the spectrum of the isospectral path B_t given by (5.10). In particular,

$$\text{dist}(\omega_t, \Omega_t) = \pi, \quad t \in \left[0, \frac{\pi}{2}\right],$$

and hence

$$\frac{\pi}{2} \int_0^t \frac{\|\overline{\dot{B}_\tau}\|}{\text{dist}(\omega_\tau, \Omega_\tau)} d\tau = \frac{\pi}{2} \int_0^t \frac{2}{\pi} d\tau = t, \quad t \in \left[0, \frac{\pi}{2}\right].$$

To complete the proof of (5.14) it suffices to show that

$$\arcsin(\|P_t - P_0\|) = t, \quad t \in \left[0, \frac{\pi}{2}\right].$$

We will prove a slightly more general result that states that the path of the orthogonal projections $[0, \frac{\pi}{2}] \ni t \rightarrow P_t$ is a geodesic. That is,

$$\arcsin(\|P_t - P_0\|) = \int_0^t \|\dot{P}_\tau\| d\tau = t, \quad t \in \left[0, \frac{\pi}{2}\right].$$

Introduce the notation $P = P_0$. From the definition (5.12) of the sets ω_t it follows that P is the orthogonal projection onto the closure of $\text{span}_{k \in \mathbb{Z}}\{e^{i2kx}\}$, the space generated by the “even harmonics”. From (5.10) and (5.13) it follows that

$$(5.16) \quad P_t = U_t P U_t^*, \quad t \in \left[0, \frac{\pi}{2}\right],$$

where the family of unitary operators U_t is given by (5.11).

First, we prove the inequality

$$(5.17) \quad \|P_t - P\| \geq \sin t, \quad t \in \left[0, \frac{\pi}{2}\right],$$

We proceed as follows.

One observes that

$$\begin{aligned} \|(U_t P U_t^* - P)\| &= \|(U_t P U_t^* - P)U_t\| \\ &\geq \|(U_t P U_t^* - P)U_t P^\perp\| = \|P U_t P^\perp\|, \quad t \in \left[0, \frac{\pi}{2}\right]. \end{aligned}$$

The operator PU_tP^\perp can easily be shown to be unitarily equivalent (up to a scalar factor) to the regularized discrete Hilbert transform H_p in $\ell^2(\mathbb{Z})$ with $p = \frac{2t+\pi}{2\pi}$,

$$(5.18) \quad PU_tP^\perp \sim -\frac{\sin 2t}{2\pi} H_p, \quad t \in \left(0, \frac{\pi}{2}\right),$$

where the symbol \sim denotes a unitary equivalence.

Recall that by the definition the regularized Hilbert transform H_p is given by the following convolution operator

$$(H_p \hat{a})_m = \sum_{n \in \mathbb{Z}} \frac{a_n}{m - n + p}, \quad \hat{a} = \{a_n\}_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}), \quad p \in (0, 1).$$

Indeed, to prove (5.18), take a $g \in \text{Ran } P^\perp$ with the Fourier series

$$g(x) = \sum_k g_k e^{i(2k+1)\pi x}.$$

Then the Fourier series of the function $PU_tP^\perp g$ is given by

$$(5.19) \quad \begin{aligned} (PU_tP^\perp g)(x) &= \sum_m \sum_k g_k \frac{1}{2} \left(\int_{-1}^1 e^{i((2(k-m)+1)\pi+2t)\tau} d\tau \right) e^{2im\pi x} \\ &= -\sum_{k,m} g_k \frac{\sin 2t}{(2(k-m)+1)\pi+2t} e^{2im\pi x} \\ &= -\frac{\sin 2t}{2\pi} \sum_m (H_p \hat{g})_k e^{2im\pi x}, \quad \hat{g} = \{g_k\}_{k \in \mathbb{Z}} \in \ell^2. \end{aligned}$$

Representation (5.19) proves the claim (5.18). In particular,

$$(5.20) \quad \|PU_tP^\perp\| = \frac{\sin 2t}{2\pi} \|H_p\|_{\ell^2(\mathbb{Z})}, \quad t \in \left(0, \frac{\pi}{2}\right).$$

Next, the symbol h_p of the convolution operator H_p can be computed explicitly and it is given by

$$h_p(x) = \frac{\pi}{\sin \pi p} e^{i\pi p(1-x)} = \sum_m \frac{e^{im\pi x}}{m+p}, \quad x \in (0, 2).$$

Hence, the norm of H_p in the space $\ell^2(\mathbb{Z})$ coincides with the ℓ^∞ -norm of the symbol h_p and therefore

$$(5.21) \quad \|H_p\|_{\ell^2(\mathbb{Z})} = \sup_{x \in [-1, 1]} \left| \frac{\pi e^{i\pi p(1-x)}}{\sin \pi p} \right| = \frac{\pi}{\sin \pi p} = \frac{\pi}{\cos t}, \quad t \in \left(0, \frac{\pi}{2}\right).$$

Combining (5.20) with (5.21) yields the lower bound (5.17), that is,

$$(5.22) \quad \|P_t - P\| \geq \|PU_tP^\perp\| = \frac{\sin 2t}{2\pi} \frac{\pi}{\cos t} = \sin t, \quad t \in \left(0, \frac{\pi}{2}\right).$$

Our next immediate goal is to prove the opposite inequality

$$(5.23) \quad \|P_t - P\| \leq \sin t, \quad t \in \left[0, \frac{\pi}{2}\right].$$

Using the result of Lemma 3.4, it is sufficient to prove that

$$(5.24) \quad \|\dot{P}_t\| = 1, \quad t \in \left(0, \frac{\pi}{2}\right).$$

In order to prove (5.24), one observes that

$$\|\dot{P}_t\| = \|U_t[2\hat{x}, P]U_t^*\| = 2\|[\hat{x}, P]\|, \quad t \geq 0,$$

with \hat{x} the multiplication operator by the independent variable,

$$(\hat{x}f)(x) = xf(x), \quad f \in L^2(-1, 1).$$

So, $\|\dot{P}_t\|$ does not depend on the parameter t . Therefore, it remains to show that

$$\|\dot{P}_0\| = 1.$$

Indeed,

$$i^{-1}[2\hat{x}, P] = iP2\hat{x}P^\perp - iP^\perp 2\hat{x}P = iP2\hat{x}P^\perp + (iP2\hat{x}P^\perp)^*$$

and hence the commutator $i^{-1}[2\hat{x}, P]$ can be represented as the following off-diagonal self-adjoint operator matrix with respect to the decomposition $L^2(-1, 1) = \text{Ran } P \oplus \text{Ran } P^\perp$

$$i^{-1}[2\hat{x}, P] = \begin{pmatrix} 0 & V \\ V^* & 0 \end{pmatrix}.$$

Here the bounded operator $V \in \mathcal{L}(\text{Ran } P^\perp, \text{Ran } P)$, is given by

$$V = i2P\hat{x}|_{\text{Ran } P^\perp}.$$

However, it follows from (5.22) that

$$\|V\| = \left\| \lim_{t \downarrow 0} \frac{P(U_t - I)P^\perp}{it} \right\| = \left\| \lim_{t \downarrow 0} \frac{PU_t P^\perp}{it} \right\| = \lim_{t \downarrow 0} \frac{\sin t}{t} = 1.$$

Hence,

$$\|[2\hat{x}, P]\| = \left\| \begin{pmatrix} 0 & V \\ V^* & 0 \end{pmatrix} \right\| = 1.$$

Combining (5.17), (5.23) and (5.24) proves that the path $[0, \frac{\pi}{2}] \ni t \rightarrow P_t$ is geodesic, and hence (5.14) holds. \square

6. NEW ESTIMATES IN THE SUBSPACE PERTURBATION PROBLEM

The main goal of this section is to apply Theorem 5.2 to the solution of the subspace perturbation problem recently discussed in [5] and [7].

Recall that if A and V are self-adjoint bounded operators and A has a spectral component ω separated from the rest of the spectrum Ω , then the spectrum of $A + V$ still consists of two separated parts, provided that $\|V\|$ is small enough. Due to the upper semicontinuity of the spectrum (cf., e.g., [4, Theorem V.4.10]) this is the case if the (in general sharp) condition $\|V\| < d/2$ with $d = \text{dist}(\omega, \Omega)$ is satisfied. Moreover, if the perturbation V is off-diagonal with respect to the decomposition $\mathcal{H} = \text{Ran } E_A(\omega) \oplus \text{Ran } E_A(\Omega)$, in [5] it is shown, that the optimal gap nonclosing condition is $\|V\| < \frac{\sqrt{3}}{2}d$ and this conditions is sharp as well.

It is now a natural question, under what (possibly stronger) condition on the norm of V the difference of the spectral projections for A and $A+V$ associated with the corresponding spectral components is a contraction with the norm less than 1.

Our first application of Theorem 5.2 treats the case of arbitrary bounded self-adjoint perturbations V .

Theorem 6.1. *Assume that A and V are bounded self-adjoint operators. Suppose that the spectrum of A has a part ω separated from the remainder of the spectrum Ω in the sense that*

$$(6.1) \quad \text{spec}(A) = \omega \cup \Omega \quad \text{and} \quad \text{dist}(\omega, \Omega) = d > 0.$$

If

$$\|V\| < \frac{\sinh(1)}{e}d,$$

then

$$(6.2) \quad \|E_A(\omega) - E_{A+V}(\mathcal{O}_{d/2}(\omega))\| \leq \sin\left(\frac{\pi}{4} \log \frac{d}{d - 2\|V\|}\right) < 1,$$

where $\mathcal{O}_{d/2}(\omega)$ denotes the open $d/2$ -neighborhood of ω .

Proof. Introduce the path

$$I = [0, 1] \ni t \mapsto B_t = A + tV$$

and set

$$(6.3) \quad \omega_t := \text{spec}(B_t) \cap \mathcal{O}_{d/2}(\omega) \quad \text{and} \quad \Omega_t := \text{spec}(B_t) \cap \mathcal{O}_{d/2}(\Omega), \quad t \in I.$$

Since

$$\|V\| < \frac{\sinh(1)}{e} d < \frac{d}{2},$$

by [4, Theorem V.4.10] the families $\{\omega_t\}_{t \in I}$ and $\{\Omega_t\}_{t \in I}$ are separated with the distance function $d(t)$ satisfying the estimate

$$d(t) := \text{dist}(\omega_t, \Omega_t) \geq d - 2t\|V\| > 0, \quad t \in I.$$

Moreover, these families are also upper semicontinuous on I (cf., [4, Theorem IV.3.16]). Since the path $I \ni t \mapsto B_t$ is obviously a C^1 -smooth path (in fact, it is real analytic), from Theorem 5.2 it follows that

$$(6.4) \quad \arcsin(\|E_A(\omega) - E_{B_t}(\omega_t)\|) \leq \frac{\pi}{2} \int_0^t \frac{\|V\|}{d - 2\tau\|V\|} d\tau, \quad t \in [0, 1).$$

Observing that

$$(6.5) \quad \int_0^1 \frac{\|V\|}{d - 2\tau\|V\|} d\tau = \frac{1}{2} \log \left(\frac{d}{d - 2\|V\|} \right)$$

and

$$\frac{1}{2} \log \left(\frac{d}{d - 2\|V\|} \right) < \frac{1}{2} \log \left(\frac{1}{1 - 2\frac{\sinh(1)}{e}} \right) = 1,$$

from (6.4) (by going to the limit when t approaches 1) one gets the estimate

$$\|E_A(\omega) - E_{B_1}(\omega_1)\| = \sin \left(\frac{\pi}{4} \log \frac{d}{d - 2\|V\|} \right).$$

To complete the proof it remains to observe that $B_1 = A + V$ and that

$$E_{A+V}(\mathcal{O}_{d/2}(\omega)) = E_{B_1}(\omega_1)$$

as it follows from (6.3). □

Our second application of Theorem 5.2 concerns the case of off-diagonal perturbations where the corresponding spectral shift is rather specific. In that case, the additional knowledge about the behavior of the spectral parts from [5] gives rise to a stronger estimate compared to that in Theorem 6.1.

Theorem 6.2. *Assume the hypothesis of Theorem 6.1. Suppose, in addition, that V is off-diagonal with respect to the orthogonal decomposition $\mathcal{H} = \text{Ran } E_A(\omega) \oplus \text{Ran } E_A(\Omega)$, that is,*

$$E_A(\omega)V E_A(\omega) = E_A(\Omega)V E_A(\Omega) = 0.$$

Suppose further that

$$(6.6) \quad \|V\| < \mathfrak{s}d,$$

where \mathfrak{s} is the unique root of the equation

$$(6.7) \quad \int_0^{\mathfrak{s}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}} = 1.$$

Then

$$(6.8) \quad \|E_A(\omega) - E_{A+V}(\mathcal{O}_{d/2}(\omega))\| \leq \sin \left(\frac{\pi}{2} \int_0^{\|V\|/d} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}} \right) < 1.$$

Proof. As in the proof of Theorem 6.1, introduce the path $I = [0, 1] \ni t \mapsto B_t = A + tV$ and the sets

$$(6.9) \quad \omega_t := \text{spec}(B_t) \cap \mathcal{O}_{d/2}(\omega) \quad \text{and} \quad \Omega_t := \text{spec}(B_t) \cap \mathcal{O}_{d/2}(\Omega), \quad t \in I.$$

Since the improper integral

$$\int_0^{\frac{\sqrt{3}}{2}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}} = \infty$$

diverges, the root \mathfrak{s} of (6.7) is well-defined and less than $\frac{\sqrt{3}}{2}$ and hence

$$(6.10) \quad \|V\| < \frac{\sqrt{3}}{2}d,$$

as it follows from (6.6).

By [5, Theorem 1.3], under the condition (6.10) it is

$$\text{spec}(B_1) \cap \mathcal{O}_{d/2}(\omega) = \text{spec}(B_1) \cap \mathcal{U}_{\delta_V}(\omega)$$

and also

$$\text{spec}(B_1) \cap \mathcal{O}_{d/2}(\Omega) = \text{spec}(B_1) \cap \mathcal{U}_{\delta_V}(\Omega),$$

where

$$\delta_V = \|V\| \tan\left(\frac{1}{2} \arctan \frac{2\|V\|}{d}\right),$$

and $\mathcal{U}_{\delta_V}(\Delta)$ denotes the closed δ_V -neighborhood of the Borel set $\Delta \subset \mathbb{R}$. Therefore

$$(6.11) \quad \omega_t = \text{spec}(B_t) \cap \mathcal{U}_{\delta_{tV}}(\omega) \quad \text{and} \quad \Omega_t = \text{spec}(B_t) \cap \mathcal{U}_{\delta_{tV}}(\Omega), \quad t \in I.$$

In particular, the families $\{\omega_t\}_{t \in I}$ and $\{\Omega_t\}_{t \in I}$ are separated with the distance function $d(t)$ satisfying the estimate

$$\begin{aligned} d(t) &:= \text{dist}(\omega_t, \Omega_t) \geq d - 2t\|V\| \tan\left(\frac{1}{2} \arctan \frac{2t\|V\|}{d}\right) \\ &= \left(2 - \sqrt{1 + 4\left(\frac{t\|V\|}{d}\right)^2}\right) d > 0, \quad t \in I. \end{aligned}$$

Using a similar argument as in the proof of Theorem 6.1 and applying Theorem 5.2, one gets the estimate

$$\begin{aligned} \arcsin(\|E_A(\omega) - E_{B_1}(\omega_1)\|) &\leq \frac{\pi}{2} \frac{\|V\|}{d} \int_0^1 \frac{d\tau}{2 - \sqrt{1 + 4\left(\frac{\|V\|\tau}{d}\right)^2}} \\ &= \frac{\pi}{2} \int_0^{\frac{\|V\|}{d}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}}. \end{aligned}$$

By (6.6) and (6.7) it is

$$\int_0^{\frac{\|V\|}{d}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}} < \int_0^{\mathfrak{s}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}} = 1,$$

and one arrives at the estimate

$$\|E_A(\omega) - E_{A+V}(\mathcal{O}_{d/2}(\omega))\| = \|E_A(\omega) - E_{B_1}(\omega_1)\| \leq \sin\left(\frac{\pi}{2} \int_0^{\frac{\|V\|}{d}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}}\right),$$

which proves (6.8). \square

Remark 6.3. *In the situation of Theorem 6.1, the previously known estimate obtained in [7] under the assumption*

$$\|V\| < \frac{2}{2 + \pi}d$$

has the form

$$(6.12) \quad \|\mathbf{E}_A(\omega) - \mathbf{E}_{A+V}(\mathcal{O}_{d/2}(\omega))\| \leq \frac{\pi}{2} \frac{\|V\|}{d - \|V\|}$$

and one can show (see Appendix A) that the estimate (6.2) is stronger than (6.12), i.e.

$$(6.13) \quad \sin\left(\frac{\pi}{4} \log \frac{d}{d - 2\|V\|}\right) < \frac{\pi}{2} \frac{\|V\|}{d - \|V\|} \quad \text{whenever } 0 < \|V\| < \frac{d}{2}.$$

In the off-diagonal case of Theorem 6.2, the previously known estimate obtained in [5] under the assumption

$$(6.14) \quad \|V\| < \frac{3\pi - \sqrt{\pi^2 + 32}}{\pi^2 - 4}d$$

has the form

$$(6.15) \quad \|\mathbf{E}_A(\omega) - \mathbf{E}_{A+V}(\mathcal{O}_{d/2}(\omega))\| \leq \frac{\pi}{2} \frac{\|V\|}{d - \|V\| \tan\left(\frac{1}{2} \arctan \frac{2\|V\|}{d}\right)}.$$

Recall that the critical constant $c_\pi = \frac{3\pi - \sqrt{\pi^2 + 32}}{\pi^2 - 4}$ in (6.14) was chosen to be the only positive root of the equation

$$\frac{\pi}{2} \frac{x}{1 - x \tan\left(\frac{1}{2} \arctan 2x\right)} = 1$$

and therefore in the critical case $\|V\| = c_\pi d$ the right hand side of (6.15) turns out to be 1 which means that the bound (6.15) is not informative for the range of perturbations $\|V\|$ such that $\|V\| \geq c_\pi d$.

Note, that the identity

$$\frac{1}{2 - \sqrt{1 + 4\tau^2}} = \frac{1}{1 - 2\tau \tan\left(\frac{1}{2} \arctan 2\tau\right)}$$

holds for all $0 \leq \tau < \frac{\sqrt{3}}{2}$. Combined with the inequality

$$(6.16) \quad \sin\left(\frac{\pi}{2} \int_0^t \frac{d\tau}{1 - 2\tau \tan\left(\frac{1}{2} \arctan 2\tau\right)}\right) < \frac{\pi}{2} \frac{t}{1 - t \tan\left(\frac{1}{2} \arctan 2t\right)},$$

$$0 \leq t < \frac{\sqrt{3}}{2},$$

which is proven in Appendix B, this means that the right hand side of (6.8) is less than the right hand side of (6.15) and therefore the bound (6.8) is stronger than the previously known estimate (6.15). In particular, since the critical constant c_π was defined so that the right hand side of (6.15) equals 1 for $\|V\| = c_\pi d$, it follows immediately from (6.16) that $c_\pi < \mathfrak{s}$. Numerical calculations suggest that the exact value of \mathfrak{s} satisfies the two-sided estimate

$$0.67598931 < \mathfrak{s} < 0.67598932.$$

APPENDIX A. PROOF OF INEQUALITY (6.13)

We write $\|V\| = \alpha d$ and substitute $x = \frac{\alpha}{1-\alpha}$, $0 < x < 1$. With $\frac{1}{1-2\alpha} = \frac{1+x}{1-x}$ inequality (6.13) becomes

$$(A.1) \quad \sin\left(\frac{\pi}{4} \log \frac{1+x}{1-x}\right) < \frac{\pi}{2}x, \quad 0 < x < 1.$$

Since the left-hand side of (A.1) is not greater than 1, we may assume $x \leq \frac{2}{\pi}$. In that case, (A.1) can be rewritten as

$$(A.2) \quad \frac{\pi}{4} \log \frac{1+x}{1-x} < \arcsin\left(\frac{\pi}{2}x\right), \quad 0 < x \leq \frac{2}{\pi}.$$

It suffices to show that the corresponding inequality holds for the derivatives of both sides. Differentiating the left-hand side gives

$$\frac{d}{dx} \frac{\pi}{4} \log \frac{1+x}{1-x} = \frac{\pi}{2} \cdot \frac{1}{(1-x^2)}$$

and differentiating the right-hand side gives

$$\frac{d}{dx} \arcsin\left(\frac{\pi}{2}x\right) = \frac{\pi}{\sqrt{4-\pi^2x^2}}.$$

Therefore, we have to show that

$$(A.3) \quad \frac{\pi}{2} \cdot \frac{1}{(1-x^2)} < \frac{\pi}{\sqrt{4-\pi^2x^2}}$$

holds for all $0 < x \leq \frac{2}{\pi}$. Taking the square, we can rewrite (A.3) as

$$4 - \pi^2x^2 < 4(1-x^2)^2$$

which is equivalent to

$$0 < (\pi^2 - 2)x^2 + x^4, \quad 0 < x \leq \frac{2}{\pi}.$$

Since $\pi^2 > 2$, this is obviously true, so (A.3) holds for all $0 < x \leq \frac{2}{\pi}$, which proves (6.13).

APPENDIX B. PROOF OF INEQUALITY (6.16)

First, we remark that

$$1 - 2x \tan\left(\frac{1}{2} \arctan 2x\right) = 2 - \sqrt{1+4x^2}$$

and that

$$1 - x \tan\left(\frac{1}{2} \arctan 2x\right) = \frac{3}{2} - \frac{\sqrt{1+4x^2}}{2}$$

for $0 \leq x < \frac{\sqrt{3}}{2}$, and thus the inequality (6.16) can be rewritten as

$$(B.1) \quad \sin\left(\frac{\pi}{2} \int_0^t \frac{d\tau}{2 - \sqrt{1+4\tau^2}}\right) < \frac{\pi}{2} \frac{2t}{3 - \sqrt{1+4t^2}}.$$

It is sufficient to prove the corresponding inequality for the derivatives that, after elementary computations, can be written as

$$(B.2) \quad \cos\left(\frac{\pi}{2} \int_0^t \frac{d\tau}{2 - \sqrt{1+4\tau^2}}\right) < 2 \left[\frac{2 - \sqrt{1+4t^2}}{3 - \sqrt{1+4t^2}} + \frac{4t^2(2 - \sqrt{1+4t^2})}{(3 - \sqrt{1+4t^2})^2 \sqrt{1+4t^2}} \right].$$

After the change of variables $x = \sqrt{1 + 4t^2}$, so that $t^2 = \frac{x^2 - 1}{4}$, the desired estimate may be rewritten as

$$\cos\left(\frac{\pi}{2} \int_0^{\frac{\sqrt{x^2-1}}{2}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}}\right) < F(x), \quad 1 < x < 2,$$

where the function F is given by

$$F(x) = \frac{2(3x - 1)(2 - x)}{(3 - x)^2 x}.$$

Denote by \mathfrak{x} the first root of the equation

$$\cos\left(\frac{\pi}{2} \int_0^{\frac{\sqrt{x^2-1}}{2}} \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}}\right) = F(x)$$

that is greater than 1.

An elementary analysis shows that the equation $1 = F(x)$ has three roots

$$x_1 = -\frac{\sqrt{17} + 1}{2}, \quad x_2 = 1 \quad \text{and} \quad x_3 = \frac{\sqrt{17} - 1}{2},$$

and that

$$1 < F(x) \quad \text{on} \quad (1, x_3).$$

Therefore,

$$\mathfrak{x} > x_3 = \frac{\sqrt{17} - 1}{2}$$

and thus, we have proven the inequality

$$(B.3) \quad \sin\left(\frac{\pi}{2} \int_0^t \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}}\right) < \frac{\pi}{2} \frac{2t}{3 - \sqrt{1 + 4t^2}} \quad \text{on the interval} \quad (0, \mathfrak{t}),$$

where \mathfrak{t} is given by

$$\mathfrak{t} = \frac{\sqrt{x_3^2 - 1}}{2} = \frac{\sqrt{(\sqrt{17} - 1)^2 - 4}}{4} \approx 0.599 \dots$$

Next, we show that the right hand side of (B.3) at the point $t = \mathfrak{t}$ is greater than 1. Indeed,

$$(B.4) \quad \begin{aligned} \frac{\pi}{2} \frac{2\mathfrak{t}}{3 - \sqrt{1 + 4\mathfrak{t}^2}} &= \frac{\pi}{2} \frac{2 \frac{\sqrt{(\sqrt{17}-1)^2-4}}{4}}{3 - \frac{\sqrt{17}-1}{2}} > \pi \frac{\frac{\sqrt{(\sqrt{16}-1)^2-4}}{4}}{3 - \frac{\sqrt{17}-1}{2}} = \pi \frac{\sqrt{5}}{14 - 2\sqrt{17}} \\ &> \pi \frac{\sqrt{5}}{5} = \frac{\pi}{\sqrt{5}} > 1, \end{aligned}$$

where we used the obvious inequality $2\sqrt{17} < 9$.

From that it follows, that the right hand side of (6.16) is greater than 1 for $t \geq \mathfrak{t}$ and one concludes that

$$(B.5) \quad \sin\left(\frac{\pi}{2} \int_0^t \frac{d\tau}{2 - \sqrt{1 + 4\tau^2}}\right) < \frac{\pi}{2} \frac{2t}{3 - \sqrt{1 + 4t^2}} \quad \text{for all } t \in \left(0, \frac{\sqrt{3}}{2}\right).$$

We remark that at $t = \frac{\sqrt{3}}{2}$ the right hand side of (B.5) blows up.

APPENDIX C. ALTERNATIVE PROOF OF THEOREM 5.2

In this appendix we present an alternative proof of Theorem 5.2.

We make the following preparations.

Lemma C.1. *Under the assumptions of Theorem 5.2 we define $d: I \times I \rightarrow \mathbb{R}_0^+$ by*

$$d(t, s) := \text{dist}(\omega_t, \Omega_s).$$

Then for every $t \in I$ one has

$$\lim_{s \rightarrow t} d(t, s) = \lim_{s \rightarrow t} d(s, t) = d(t, t).$$

Furthermore, the function given by $t \mapsto d(t, t)$, $t \in I$, is continuous.

Proof. Let $t \in I$ and let $0 < \varepsilon < \frac{1}{4}d(t, t)$ be arbitrary. Since the families $\{\omega_s\}_{s \in I}$ and $\{\Omega_s\}_{s \in I}$ are upper semicontinuous, we can choose $\delta > 0$ such that

$$\omega_s \subset \mathcal{O}_\varepsilon(\omega_t), \quad \Omega_s \subset \mathcal{O}_\varepsilon(\Omega_t), \quad s \in I, \quad |s - t| < \delta,$$

as well as

$$\omega_t \subset \mathcal{O}_\varepsilon(\omega_s), \quad \Omega_t \subset \mathcal{O}_\varepsilon(\Omega_s), \quad s \in I, \quad |s - t| < \delta,$$

where $\mathcal{O}_\varepsilon(\Delta)$ denotes the open ε -neighborhood of $\Delta \subset \mathbb{R}$. From that one obtains

$$d(t, t) - 2\varepsilon \leq d(s, s) \leq d(t, t) + 2\varepsilon$$

and

$$d(t, t) - \varepsilon \leq d(t, s) \leq d(t, s) + \varepsilon$$

for all $s \in I$ such that $|s - t| < \delta$. The same is true for $d(s, t)$ instead of $d(t, s)$, which completes the proof. \square

Proposition C.2 ([8]). *Let A and B be bounded self-adjoint operators and ω and Ω two Borel sets on the real line. Then*

$$\text{dist}(\omega, \Omega) \|E_A(\omega)E_B(\Omega)\| \leq \frac{\pi}{2} \|A - B\|.$$

Moreover, if the convex hull of the set ω does not intersect the set Ω , or vice versa, then

$$\text{dist}(\omega, \Omega) \|E_A(\omega)E_B(\Omega)\| \leq \|A - B\|.$$

Now we are able to prove Theorem 5.2.

Proof of Theorem 5.2. By Proposition C.2,

$$(C.1) \quad \text{dist}(\omega_t, \Omega_s) \|P_t P_s^\perp\| \leq \frac{\pi}{2} \|B_t - B_s\|, \quad s, t \in I,$$

and

$$(C.2) \quad \text{dist}(\omega_s, \Omega_t) \|P_t^\perp P_s\| \leq \frac{\pi}{2} \|B_t - B_s\|, \quad s, t \in I.$$

Since

$$\|P_s - P_t\| = \max \left\{ \|P_t P_s^\perp\|, \|P_t^\perp P_s\| \right\},$$

from (C.1) and (C.2) it follows that

$$\min \{ \text{dist}(\omega_t, \Omega_s), \text{dist}(\omega_s, \Omega_t) \} \|P_s - P_t\| \leq \frac{\pi}{2} \|B_s - B_t\|, \quad s, t \in I.$$

Dividing both sides of this inequality by $|s - t|$ and letting s approach t , one obtains the bound

$$\text{dist}(\omega_t, \Omega_t) \|\dot{P}_t\| \leq \frac{\pi}{2} \|\dot{B}_t\|, \quad t \in I,$$

where we have used Lemma C.1 and the smoothness of the path $I \ni t \mapsto P_t$ (cf. Appendix D).

Since $\text{dist}(\omega_t, \Omega_t) > 0$ for all $t \in I$ by hypothesis, one obtains that

$$\|\dot{P}_t\| \leq \frac{\pi}{2} \frac{\|\dot{B}_t\|}{\text{dist}(\omega_t, \Omega_t)}, \quad t \in I,$$

and then applying Lemma 3.4 completes the proof. \square

APPENDIX D. PROOF OF THE SMOOTHNESS OF THE SPECTRAL PROJECTIONS

The proof of the smoothness of the path of projections P_t required for the alternative proof of Theorem 5.2 in Appendix C is essentially the same as the one presented in [4, Theorem II.5.4] for the continuous case.

Lemma D.1. *Under the assumptions of Theorem 5.2, $I \ni t \mapsto P_t$ is a C^1 -smooth path.*

Proof. Let $t \in I$ and $\varepsilon = \frac{1}{4}\text{dist}(\omega_t, \Omega_t) > 0$. Due to the fact, that the families $\{\omega_s\}_{s \in I}$ and $\{\Omega_s\}_{s \in I}$ are upper semicontinuous, there is a $\delta > 0$ such that

$$(D.1) \quad \omega_s \subset \mathcal{O}_{\varepsilon/2}(\omega_t) \text{ and } \Omega_s \subset \mathcal{O}_{\varepsilon/2}(\Omega_t) \quad \text{for all } s \in I, |s - t| < \delta,$$

where $\mathcal{O}_\varepsilon(\Delta)$ denotes the open ε -neighborhood of $\Delta \subset \mathbb{R}$.

In particular, $\mathcal{O}_\varepsilon(\omega_t) \setminus \mathcal{O}_{\varepsilon/2}(\omega_t)$ lies in the resolvent set of B_s for all $s \in I$, $|s - t| < \delta$. Therefore, there exists a finite number of rectifiable, simple closed positive orientated curves belonging to $\mathbb{C} \setminus \text{spec}(B_s)$ for all $s \in I$, $|s - t| < \delta$, such that ω_s is contained in the union of their interiors and Ω_s lies in the union of their exteriors. Let Γ denote the union of these curves. As in [4, (III.6.19)], P_s has the representation

$$P_s = \frac{1}{2\pi i} \int_{\Gamma} R_s(\zeta) d\zeta, \quad R_s(\zeta) := (\zeta I_{\mathcal{H}} - B_s)^{-1}, \quad s \in I, |s - t| < \delta.$$

Since $B_s - B_t = (\zeta I_{\mathcal{H}} - B_t) - (\zeta I_{\mathcal{H}} - B_s)$, it is

$$R_s(\zeta) - R_t(\zeta) = R_t(\zeta)(B_s - B_t)R_s(\zeta), \quad s \in I, |s - t| < \delta.$$

Furthermore, for all $\zeta \in \Gamma \subset \mathbb{C} \setminus \text{spec}(B_s)$ the equation

$$(D.2) \quad \|R_s(\zeta)\| = \frac{1}{\text{dist}(\zeta, \text{spec}(B_s))}$$

holds (cf. [4, (V.3.16)]). Due to (D.1), this implies that $\|R_s(\zeta)\|$ is uniformly bounded for $\zeta \in \Gamma$ and $s \in I$, $|s - t| < \delta$, from which one concludes, that $\frac{R_s(\zeta) - R_t(\zeta)}{s - t}$ converges uniformly to $R_t(\zeta)\dot{B}_t R_t(\zeta)$ for $\zeta \in \Gamma$ as s goes to t . This shows that

$$\dot{P}_t = \lim_{s \rightarrow t} \frac{P_s - P_t}{s - t} = \frac{1}{2\pi i} \int_{\Gamma} R_t(\zeta) \dot{B}_t R_t(\zeta) d\zeta$$

exists. By a similar argument, one concludes that $I \ni t \mapsto \dot{P}_t$ is continuous and, therefore, $I \ni t \mapsto P_t$ is C^1 -smooth. \square

REFERENCES

- [1] N. I. Achieser, I. M. Glasmann, *Theory of Linear Operators in Hilbert Space*, Dover Publications, New York, 1993.
- [2] M. R. Bridson, A. Haefliger, *Metric spaces of non-positive curvature*, Springer, Berlin, 1999.
- [3] P. R. Halmos, *Two subspaces*, Trans. Amer. Math. Soc. **144** (1969), 381–389.
- [4] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1966.
- [5] V. Kostykin, K. A. Makarov, A. K. Motovilov, *Perturbation of spectra and spectral subspaces*, **359** (2007), 77–89.
- [6] V. Kostykin, K. A. Makarov, A. K. Motovilov, *Existence and Uniqueness of Solutions to the Operator Riccati Equation. A Geometric Approach*, Contemporary Mathematics Vol. 327, Amer. Math. Soc., 2003, p. 181 – 198.
- [7] V. Kostykin, K. A. Makarov, A. K. Motovilov, *On a subspace perturbation problem*, Proc. Amer. Math. Soc. **131** (2003), 3469–3476.

- [8] R. McEachin, *Closing the gap in a subspace perturbation bound*, Linear Algebra Appl. bf 180 (1993), 7–15.
- [9] B. Sz.-Nagy, *Über die Ungleichung von H. Bohr.*, Math. Nachr. **9** (1953), 255–259.

K. A. MAKAROV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MISSOURI, COLUMBIA, MO 65211, USA

E-mail address: makarovk@missouri.edu

A. SEELMANN, FB 08 - INSTITUT FÜR MATHEMATIK, JOHANNES GUTENBERG-UNIVERSITÄT MAINZ, STAUDINGER WEG 9, D-55099 MAINZ, GERMANY

E-mail address: seelmann@mathematik.uni-mainz.de